

**TADG-15: AN EXTRACELLULAR SERINE PROTEASE  
OVEREXPRESSED IN CARCINOMAS**

5

10

**BACKGROUND OF THE INVENTION**

Cross-Reference to Related Application

This application is a continuation-in-part of USSN 09/027,337, filed February 20, 1998 and thereby claims the benefit 15 of priority under 35 USC §120.

Field of the Invention

The present invention relates generally to the fields of cellular biology and the diagnosis of neoplastic disease. More 20 specifically, the present invention relates to an extracellular serine

protease, termed tumor antigen-derived gene 15 (TADG-15), which is overexpressed in carcinomas.

#### Description of the Related Art

5                   Extracellular proteases have been directly associated with tumor growth, shedding of tumor cells and invasion of target organs. Individual classes of proteases are involved in, but not limited to, (a) digestion of stroma surrounding the initial tumor area, (b) digestion of the cellular adhesion molecules to allow dissociation of tumor cells; 10 and (c) invasion of the basement membrane for metastatic growth and activation of both tumor growth factors and angiogenic factors.

In the process of cancer progression and invasion, proteases mediate specific proteolysis and contribute to the removal of extracellular matrix components surrounding tumor cells, the 15 digestion of intercellular adhesion molecules to allow dissociation of malignant cells and the activation of many growth and angiogenic factors.<sup>1-3</sup> Depending on the nature of their catalytic domain, proteases are classified into four families: serine proteases, metalloproteases, aspartic proteases and cysteine proteases.<sup>3</sup> Among 20 these proteases, the metalloproteases have been well studied in relation to tumor growth and progression, and they are known to be

capable of degrading the extracellular matrix, thereby enhancing the invasive potential of malignant cells.<sup>1,4,5</sup> For serine proteases, previous studies have demonstrated an increased production of plasminogen activator in tumor cells and a positive correlation 5 between plasminogen activator activity and aggressiveness of cancer.<sup>6,7</sup> Prostate specific antigen (a serine protease) has also been widely used as an indicator of abnormal prostate growth.<sup>8</sup> More recently, several other serine proteases have been reported, *viz.* hepsin and the stratum corneum chymotryptic enzyme (SCCE), which 10 are overexpressed in ovarian cancer and which may contribute to malignant progression by increasing the extracellular lytic activity of these tumor cells.<sup>9</sup>

The prior art is deficient in the lack of effective means of screening to identify proteases overexpressed in carcinoma. The 15 present invention fulfills this longstanding need and desire in the art.

## SUMMARY OF THE INVENTION

20 The present invention discloses a screening program to identify proteases overexpressed in carcinoma by examining PCR

products amplified using differential display in early stage tumors and metastatic tumors compared to that of normal tissues. The approach herein to identify candidate genes overexpressed in tumor cells has been to utilize the well conserved domains surrounding the 5 triad of amino acids (His-Asp-Ser) prototypical of the catalytic domain of serine proteases. Herein, evidence is presented for a unique form of serine protease not previously described in the literature which is highly expressed in ovarian carcinomas. Through the screening approach using differential PCR amplification of 10 normal, low malignant potential and overt carcinomas, a PCR product present only in carcinoma was subcloned and sequenced, and was found to have a catalytic domain which was consistent with the serine protease family. Reported herein is the complete cloning and sequencing of this transcript and evidence for its expression in 15 ovarian tumor cells.

In one embodiment of the present invention, there is provided a DNA encoding a tumor antigen-derived gene (TADG-15) protein, selected from the following: (a) an isolated DNA which encodes a TADG-15 protein; (b) an isolated DNA which hybridizes 20 under high stringency conditions to the isolated DNA of (a) above and which encodes a TADG-15 protein; and (c) an isolated DNA differing

from the isolated DNAs of (a) and (b) above in codon sequence due to the degeneracy of the genetic code, and which encodes a TADG-15 protein. The embodiment further includes a vector comprising the TADG-15 DNA and regulatory elements necessary for expression of 5 the DNA in a cell. Additionally embodied is a vector in which the TADG-15 DNA is positioned in reverse orientation relative to the regulatory elements such that TADG-15 antisense mRNA is produced.

In another embodiment of the present invention, there is provided an isolated and purified TADG-15 protein coded for by DNA 10 selected from the following: (a) an isolated DNA which encodes a TADG-15 protein; (b) an isolated DNA which hybridizes under high stringency conditions to isolated DNA of (a) above and which encodes a TADG-15 protein; and (c) an isolated DNA differing from the isolated DNAs of (a) and (b) above in codon sequence due to the 15 degeneracy of the genetic code, and which encodes a TADG-15 protein.

In yet another embodiment of the present invention, there is provided a method for detecting TADG-15 mRNA in a sample, comprising the steps of: (a) contacting a sample with a probe 20 which is specific for TADG-15; and (b) detecting binding of the probe to TADG-15 mRNA in the sample. In still yet another embodiment of

the present invention, there is provided a kit for detecting TADG-15 mRNA, comprising: an oligonucleotide probe specific for TADG-15. A label for detection is further embodied in the kit.

The present invention additionally embodies a method of 5 detecting TADG-15 protein in a sample, comprising the steps of: (a) contacting a sample with an antibody which is specific for TADG-15 or a fragment thereof; and (b) detecting binding of the antibody to TADG-15 protein in the sample. Similarly, the present invention also embodies a kit for detecting TADG-15 protein, comprising: an 10 antibody specific for TADG-15 protein or a fragment thereof. Means for detection of the antibody is further embodied in the kit.

In another embodiment, the present invention provides an antibody specific for the TADG-15 protein or a fragment thereof.

In yet another embodiment, the present invention 15 provides a method of screening for compounds that inhibit TADG-15, comprising the steps of: (a) contacting a sample comprising TADG-15 protein with a compound; and (b) assaying for TADG-15 protease activity. Typically, a decrease in the TADG-15 protease activity in the presence of the compound relative to TADG-15 protease activity 20 in the absence of the compound is indicative of a compound that inhibits TADG-15.

In still yet another embodiment of the present invention, there is provided a method of inhibiting expression of TADG-15 in a cell, comprising the step of: (a) introducing a vector into a cell, whereupon expression of the vector produces TADG-15 antisense mRNA in the cell which hybridizes to endogenous TADG-15 mRNA, thereby inhibiting expression of TADG-15 in the cell.

Further embodied by the present invention, there is provided a method of inhibiting a TADG-15 protein in a cell, comprising the step of: (a) introducing an antibody specific for a TADG-15 protein or a fragment thereof into a cell, whereupon binding of the antibody to the TADG-15 protein inhibits the TADG-15 protein.

In an embodiment of the present invention, there is provided a method of targeted therapy to an individual, comprising the step of: (a) administering a compound containing a targeting moiety and a therapeutic moiety to an individual, wherein the targeting moiety is specific for TADG-15.

In an embodiment of the present invention, there is provided a method of diagnosing cancer in an individual, comprising the steps of: (a) obtaining a biological sample from an individual; and (b) detecting TADG-15 in the sample, wherein the presence of TADG-

15 in the sample is indicative of the presence of carcinoma in the individual and the absence of TADG-15 in the sample is indicative of the absence of carcinoma in the individual.

In another embodiment of the present invention, there is  
5 provided a method of vaccinating an individual against TADG-15, comprising the steps of: (a) inoculating an individual with a TADG-15 protein or fragment thereof that lacks TADG-15 protease activity, wherein the inoculation with the TADG-15 protein or fragment thereof elicits an immune response in the individual, thereby  
10 vaccinating the individual against TADG-15.

In an embodiment of the present invention, there is provided a method of producing immune-activated cells directed toward TADG-15, comprising the steps of: exposing dendritic cells to a TADG-15 protein or fragment thereof that lacks TADG-15 protease activity, wherein the exposure to said TADG-15 protein or fragment thereof activates the dendritic cells, thereby producing immune-activated cells directed toward TADG-15.  
15

In another embodiment of the present invention, there is provided an immunogenic composition, comprising an immunogenic fragment of a TADG-15 protein and an appropriate adjuvant.  
20

Other and further aspects, features, and advantages of the present invention will be apparent from the following description of the presently preferred embodiments of the invention given for the purpose of disclosure.

5

#### BRIEF DESCRIPTION OF THE DRAWINGS

10 So that the matter in which the above-recited features, advantages and objects of the invention, as well as others which will become clear, are attained and can be understood in detail, more particular descriptions of the invention briefly summarized above may be had by reference to certain embodiments thereof which are 15 illustrated in the appended drawings. These drawings form a part of the specification. It is to be noted, however, that the appended drawings illustrate preferred embodiments of the invention and therefore are not to be considered limiting in their scope.

Figure 1 shows a comparison of the serine protease 20 catalytic domain of TADG-15 with Hepsin (Heps, SEQ ID No. 3), SCCE (SEQ ID No. 4), Trypsin (Try, SEQ ID No. 5), Chymotrypsin (Chymb,

SEQ ID No. 6), Factor 7 (Fac7, SEQ ID No. 7) and Tissue plasminogen activator (Tpa, SEQ ID No. 8). The asterisks indicate conserved amino acids of catalytic triad.

Figure 2 shows the nucleotide sequence of the TADG-15 cDNA and the derived amino acid sequence of the TADG-15 protein. The putative start codon is located at nucleotides 23-25. The potential transmembrane sequence is underlined. Possible N-linked glycosylation sites are indicated by a broken line. The asterisks indicate conserved cysteine residues of CUB domain. The SDE-motifs of the LDL receptor ligand binding repeat-like domain are boxed. The arrow shows the arginine-valine bond cleaved upon activation. The conserved amino acids of the catalytic triad; histidine, aspartic acid and serine residues are circled.

Figure 3 shows the amino acid sequence of the TADG-15 protease, including functional sites and domains.

Figure 4 shows a diagram of the TADG-15 protein. 1; cytoplasmic domain, (aa #1-54), 2; transmembrane domain (aa #55-57), 3; extracellular domain (aa #78-213), 4-5; CUB repeat (aa #214-447), 6-9; LDL receptor ligand binding repeat (class A motif) like domain (aa #453-602), 10; serine protease (aa #615-855).

Figure 5 shows Northern blot analysis of TADG-15 mRNA expression in normal ovary, ovarian carcinomas, carcinoma cell lines, normal fetal tissues and normal adult tissues. A single intense transcript of the TADG-15 was observed in every sub-type of carcinoma and the two ovarian carcinoma cell lines, SW626 and CAOV3, whereas no visible band was detected in normal ovary or the two breast cancer cell lines. In normal fetal tissues, fetal kidney showed increased transcript and faint expression was detected in fetal lung. In normal adult tissues, the TADG-15 was detected in colon with low expression in small intestine and prostate.

Figure 6A shows quantitative PCR analysis of TADG-15 expression. Expression levels of TADG-15 relative to  $\beta$ -tubulin are significantly elevated in all LMP tumors and carcinomas compared to that of normal ovaries. m; mucinous, s; serous. Figure 6B shows the ratio of TADG-15 expression to expression of  $\beta$ -tubulin in normal ovary, LMP tumor and ovarian carcinoma. TADG-15 mRNA expression levels were significantly elevated in both LMP tumor (\*;  $p<0.001$ ) and carcinoma (\*\*;  $p<0.0001$ ) compared to that in normal ovary. All 10 samples of normal ovary showed a low level of TADG-15 expression.

Figure 7 shows the TADG-15 expression in tumor cell lines derived from both ovarian and breast carcinoma tissues.

Figure 8 shows the overexpression of TADG-15 in other tumor tissues.

5                   Figure 9 shows SW626 and CAOV3 cell lysates that were separated by SDS-PAGE and immunoblotted. Lanes 1 and 2 were probed with rabbit pre-immune serum as a negative control. Lanes 3 and 4 were probed with polyclonal rabbit antibody generated to a carboxy terminal peptide from TADG-15 protein sequence.

10                  Figure 10 shows that immunohistochemical staining of normal ovarian epithelium (Figure 10A) with a polyclonal antibody to a TADG-15 protease peptide shows no staining of the stroma or epithelium. However, antibody staining of carcinomas confirms the presence of TADG-15 expression in the cytoplasm of a serous low 15 malignant potential tumor (Figure 10B); a mucinous low malignant potential tumor (Figure 10C); a serous carcinoma (Figure 10D); and its presence in both the cytoplasm and cell surface of an endometrioid carcinoma (Figure 10E).

Figure 11 shows an alignment of the human TADG15 20 protein sequence with that of mouse epithin which demonstrates that the proteins are 84% similar and 81% identical over 843 amino

acids. Residues that are identical between the two proteins are indicated by a "-", while the "\*" symbol represents the TADG15 translation termination. The most significant difference between these two proteins lies in the carboxy-termini, which for epithin, 5 includes 47 amino acids that are not present in TADG15.

Figure 12 shows a nucleotide sequence comparison between TADG-15 and human SNC-19 (GeneBank Accession No. #U20428).

10

#### DETAILED DESCRIPTION OF THE INVENTION

Proteases have been implicated in the extracellular 15 modulation required for tumor growth and invasion. In an effort to categorize those proteases contributing to ovarian carcinoma progression, redundant primers directed towards conserved amino acid domains surrounding the catalytic triad of His, Asp and Ser were utilized to amplify serine proteases differentially expressed in 20 carcinomas. Using this method, a serine protease named TADG-15 (tumor antigen-derived gene 15) has been identified that is

overexpressed in ovarian tumors, TADG-15 appears to be a transmembrane multidomain serine protease. TADG-15 is highly overexpressed in ovarian tumors based on PCR, Northern blot and immunolocalization.

5                   The TADG-15 cDNA is 3147 base pairs long (SEQ ID No. 1) encoding for a 855 amino acid protein (SEQ ID No. 2). The availability of the TADG-15 gene provides numerous utilities. For example, the TADG-15 gene can be used as a diagnostic or therapeutic target in ovarian and other carcinomas, including breast, 10 prostate, lung and colon.

                  The present invention is directed to DNA encoding a tumor antigen-derived gene (TADG-15) protein, selected from the following: (a) an isolated DNA which encodes a TADG-15 protein; (b) an isolated DNA which hybridizes under high stringency conditions to 15 the isolated DNA of (a) above and which encodes a TADG-15 protein; and (c) an isolated DNA differing from the isolated DNAs of (a) and (b) above in codon sequence due to the degeneracy of the genetic code, and which encodes a TADG-15 protein. It is preferred that the DNA has the sequence shown in SEQ ID No. 1 and the TADG-15 20 protein has the amino acid sequence shown in SEQ ID No. 2.

The present invention is directed toward a vector comprising the TADG-15 DNA and regulatory elements necessary for expression of the DNA in a cell, or a vector in which the TADG-15 DNA is positioned in reverse orientation relative to the regulatory elements such that TADG-15 antisense mRNA is produced. Generally, 5 the DNA encodes a TADG-15 protein having the amino acid sequence shown in SEQ ID No. 2. The invention is also directed toward host cells transfected with either of the above-described vector(s). Representative host cells are bacterial cells, mammalian cells, plant 10 cells and insect cells. Preferably, the bacterial cell is *E. coli*.

The present invention is directed toward an isolated and purified TADG-15 protein coded for by DNA selected from the following: (a) an isolated DNA which encodes a TADG-15 protein; (b) an isolated DNA which hybridizes under high stringency conditions to 15 isolated DNA of (a) above and which encodes a TADG-15 protein; and (c) an isolated DNA differing from the isolated DNAs of (a) and (b) above in codon sequence due to the degeneracy of the genetic code, and which encodes a TADG-15 protein. Preferably, the protein has the amino acid sequence shown in SEQ ID No. 2.

20 The present invention is directed toward a method for detecting TADG-15 mRNA in a sample, comprising the steps of: (a)

contacting a sample with a probe which is specific for TADG-15; and  
(b) detecting binding of the probe to TADG-15 mRNA in the sample.

The present invention is also directed toward a method of detecting  
TADG-15 protein in a sample, comprising the steps of: (a) contacting  
5 a sample with an antibody which is specific for TADG-15 or a  
fragment thereof; and (b) detecting binding of the antibody to TADG-  
15 protein in the sample. Generally, the sample is a biological  
sample; preferably, the biological sample is from an individual; and  
typically, the individual is suspected of having cancer.

10 The present invention is directed toward a kit for  
detecting TADG-15 mRNA, comprising: an oligonucleotide probe,  
wherein the probe is specific for TADG-15. The kit may further  
comprise: a label with which to label the probe; and means for  
detecting the label. The present invention is additionally directed  
15 toward a kit for detecting TADG-15 protein, comprising: an antibody  
which is specific for TADG-15 protein or a fragment thereof. The kit  
may further comprise: means to detect the antibody.

The present invention is directed toward a antibody  
which is specific for TADG-15 protein or a fragment thereof.

20 The present invention is directed toward a method of  
screening for compounds that inhibit TADG-15, comprising the steps

of: (a) contacting a sample containing TADG-15 protein with a compound; and (b) assaying for TADG-15 protease activity. Typically, a decrease in the TADG-15 protease activity in the presence of the compound relative to TADG-15 protease activity in 5 the absence of the compound is indicative of a compound that inhibits TADG-15.

The present invention is directed toward a method of inhibiting expression of TADG-15 in a cell, comprising the step of: (a) introducing a vector expressing TADG-15 antisense mRNA into a cell, 10 which hybridizes to endogenous TADG-15 mRNA, thereby inhibiting expression of TADG-15 in the cell. Generally, the inhibition of TADG-15 expression is for treating cancer.

The present invention is directed toward a method of inhibiting a TADG-15 protein in a cell, comprising the step of: (a) 15 introducing an antibody specific for a TADG-15 protein or a fragment thereof into a cell, which inhibits the TADG-15 protein. Generally, the inhibition of the TADG-15 protein is for treating cancer.

The present invention is directed toward a method of targeted therapy to an individual, comprising the step of: (a) 20 administering a compound having a targeting moiety and a therapeutic moiety to an individual, wherein the targeting moiety is

specific for TADG-15. Representative targeting moiety are an antibody specific for TADG-15 and a ligand or ligand binding domain (e.g., CUB, LDLR, protease and extracellular) that binds TADG-15. Likewise, a representative therapeutic moiety is a radioisotope, a 5 toxin, a chemotherapeutic agent and immune stimulants. Typically, the above-described method is useful when the individual suffers from ovarian cancer, breast cancer or cancers of the prostate, lung, colon and cervix.

The present invention is directed toward a method of 10 diagnosing cancer in an individual, comprising the steps of: (a) obtaining a biological sample from an individual; and (b) detecting TADG-15 in the sample. Generally, the presence of TADG-15 in the sample is indicative of the presence of carcinoma in the individual, and the absence of TADG-15 in the sample is indicative of the 15 absence of carcinoma in the individual. Generally, the biological sample is blood, ascites fluid, urine, tears, saliva or interstitial fluid. Typical means of detecting TADG-15 are by Northern blot, Western blot; PCR, dot blot, ELIZA, radioimmunoassay, DNA chips or tumor cell labeling. This method may be useful in diagnosing cancers such as 20 ovarian, breast and other cancers in which TADG-15 is overexpressed, such as lung, prostate and colon cancers.

The present invention is also directed to an antisense oligonucleotide having the nucleotide sequence complementary to a TADG-15 mRNA sequence. The present invention is also directed to a composition comprising such an antisense oligonucleotide according 5 and a physiologically acceptable carrier therefore.

The present invention is also directed to a method of treating a neoplastic state in an individual syndrome in an individual in need of such treatment, comprising the step of administering to said individual an effective dose of an antisense oligonucleotide of.

10 Preferably, the neoplastic state is selected from the group consisting of from ovarian cancer, breast cancer, lung cancer, prostate cancer, colon cancer and other cancers in which TADG-15 is overexpressed. For such therapy, the oligonucleotides alone or in combination with other anti-neoplastic agents can be formulated for a variety of modes

15 of administration, including systemic, topical or localized administration. Techniques and formulations generally can be found in *Remington's Pharmaceutical Sciences*, Mack Publishing Co., Easton, Pa. The oligonucleotide active ingredient is generally combined with a pharmaceutically acceptable carrier such as a diluent or excipient

20 which can include fillers, extenders, binders, wetting agents, disintergrants, surface active agents or lubricants, depending on the

5 nature of the mode of administration and dosage forms. Typical dosage forms include tablets, powders, liquid preparations including suspensions, emulsions, and solutions, granules, capsules and suppositories, as well as liquid preparations for injections, including liposome preparations.

For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal and subcutaneous. For injection, the oligonucleotides of the invention are formulated in liquid solutions, preferably in physiologically compatible buffers. In addition, the oligonucleotides can be formulated in solid form and redissolved or suspended immediately prior to use. Lyophilized forms are also included. Dosages that can be used for systemic administration preferably range from about 0.01 mg/kg to 50 mg/kg administered once or twice per day.

15 However, different dosing schedules can be utilized depending on (1) the potency of an individual oligonucleotide at inhibiting the activity of its target DNA, (2) the severity or extent of the pathological disease state, or (3) the pharmacokinetic behavior of a given oligonucleotide.

20 The present invention is directed toward a method of vaccinating an individual against TADG-15 overexpression,

comprising the steps of: (a) inoculating an individual with a TADG-15 protein or fragment thereof which lacks TADG-15 protease activity. The inoculation with the TADG-15 protein or fragment thereof elicits an immune response in the individual, thereby vaccinating the  
5 individual against TADG-15. The vaccination with TADG-15 described herein is intended for an individual who has cancer, is suspected of having cancer or is at risk of getting cancer. Generally, the TADG-15 fragment useful for vaccinating an individual are 9-residue fragments up to 20-residue fragments, with preferred 9-  
10 residue fragments shown in SEQ ID Nos. 2, 19, 20, 21, 29, 39, 49, 50, 59, 79, 80, 81, 82, 83, 84, 89 and 90.

The present invention is directed toward a method of producing immune-activated cells directed toward TADG-15, comprising the steps of: exposing dendritic cells to a TADG-15 protein or fragment thereof that lacks TADG-15 protease activity, wherein exposure to the TADG-15 protein or fragment thereof activates the dendritic cells, thereby producing immune-activated cells directed toward TADG-15. Representative immune-activated cells are B-cells, T-cells and dendrites. Generally, the TADG-15 fragment is a 9-  
15 residue fragment up to a 20-residue fragment, with preferable 9-  
20 residue fragments shown in SEQ ID Nos. 2, 19, 20, 21, 29, 39, 49, 50,

59, 79, 80, 81, 82, 83, 84, 89 and 90. Preferably, the dendritic cells  
are isolated from an individual prior to exposure, and the activated  
dendritic cells reintroduced into the individual subsequent to  
exposure. Typically, the individual for which this method may apply  
5 has cancer, is suspected of having cancer or is at risk of getting  
cancer.

The present invention is directed toward an immunogenic composition, comprising an immunogenic fragment of a TADG-15 protein and an appropriate adjuvant. Generally, the fragment is a 9-  
10 residue fragment up to a 20-residue fragment, with preferred 9-  
residue fragments shown in SEQ ID Nos. 2, 19, 20, 21, 29, 39, 49, 50,  
59, 79, 80, 81, 82, 83, 84, 89 and 90.

In accordance with the present invention there may be  
employed conventional molecular biology, microbiology, and  
15 recombinant DNA techniques within the skill of the art. Such  
techniques are explained fully in the literature. See, *e.g.*, Maniatis,  
Fritsch & Sambrook, "Molecular Cloning: A Laboratory Manual  
(1982); "DNA Cloning: A Practical Approach," Volumes I and II (D.N.  
Glover ed. 1985); "Oligonucleotide Synthesis" (M.J. Gait ed. 1984);  
20 "Nucleic Acid Hybridization" (B.D. Hames & S.J. Higgins eds. 1985);  
"Transcription and Translation" (B.D. Hames & S.J. Higgins eds. 1984);

"Animal Cell Culture" (R.I. Freshney, ed. 1986); "Immobilized Cells And Enzymes" (IRL Press, 1986); B. Perbal, "A Practical Guide To Molecular Cloning" (1984). Therefore, if appearing herein, the following terms shall have the definitions set out below.

5 As used herein, the term "cDNA" shall refer to the DNA copy of the mRNA transcript of a gene.

As used herein, the term "derived amino acid sequence" shall mean the amino acid sequence determined by reading the triplet sequence of nucleotide bases in the cDNA.

10 As used herein the term "screening a library" shall refer to the process of using a labeled probe to check whether, under the appropriate conditions, there is a sequence complementary to the probe present in a particular DNA library. In addition, "screening a library" could be performed by PCR.

15 As used herein, the term "PCR" refers to the polymerase chain reaction that is the subject of U.S. Patent Nos. 4,683,195 and 4,683,202 to Mullis, as well as other improvements now known in the art.

20 The amino acid described herein are preferred to be in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the

desired functional property of immunoglobulin-binding is retained by the polypeptide. NH<sub>2</sub> refers to the free amino group present at the amino terminus of a polypeptide. COOH refers to the free carboxy group present at the carboxy terminus of a polypeptide. In 5 keeping with standard polypeptide nomenclature, *J Biol. Chem.*, 243:3552-59 (1969), abbreviations for amino acid residues are used as is customary in the art.

It should be noted that all amino-acid residue sequences are represented herein by formulae whose left and right orientation 10 is in the conventional direction of amino-terminus to carboxy-terminus. Furthermore, it should be noted that a dash at the beginning or end of an amino acid residue sequence indicates a peptide bond to a further sequence of one or more amino-acid residues.

15 A "replicon" is any genetic element (*e.g.*, plasmid, chromosome, virus) that functions as an autonomous unit of DNA replication *in vivo*; *i.e.*, capable of replication under its own control.

A "vector" is a replicon, such as plasmid, phage or cosmid, to which another DNA segment may be attached so as to bring about 20 the replication of the attached segment. A "vector" may further be

defined as a replicable nucleic acid construct, *e.g.*, a plasmid or viral nucleic acid.

A "DNA molecule" refers to the polymeric form of deoxyribonucleotides (adenine, guanine, thymine, or cytosine) in its 5 either single-stranded form or as a double-stranded helix. This term refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary forms. Thus, this term includes double-stranded DNA found, *inter alia*, in linear DNA molecules (*e.g.*, restriction fragments), viruses, plasmids, and 10 chromosomes. The structure is discussed herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (*i.e.*, the strand having a sequence homologous to the mRNA).

An expression vector is a replicable construct in which a 15 nucleic acid sequence encoding a polypeptide is operably linked to suitable control sequences capable of effecting expression of the polypeptide in a cell. The need for such control sequences will vary depending upon the cell selected and the transformation method chosen. Generally, control sequences include a transcriptional 20 promoter and/or enhancer, suitable mRNA ribosomal binding sites, and sequences which control the termination of transcription and

translation. Methods which are well known to those skilled in the art can be used to construct expression vectors containing appropriate transcriptional and translational control signals. See, for example, techniques described in Sambrook et al., 1989, *Molecular Cloning: A 5 Laboratory Manual* (2nd Ed.), Cold Spring Harbor Press, N.Y. A gene and its transcription control sequences are defined as being "operably linked" if the transcription control sequences effectively control transcription of the gene. Vectors of the invention include, but are not limited to, plasmid vectors and viral vectors. Preferred 10 viral vectors of the invention are those derived from retroviruses, adenovirus, adeno-associated virus, SV40 virus, or herpes viruses. In general, expression vectors contain promoter sequences which facilitate the efficient transcription of the inserted DNA fragment and are used in connection with the host. The expression vector typically 15 contains an origin of replication, promoter(s), terminator(s), as well as specific genes which are capable of providing phenotypic selection in transformed cells. The transformed hosts can be fermented and cultured according to means known in the art to achieve optimal cell growth.

20 An "origin of replication" refers to those DNA sequences that participate in DNA synthesis.

( (

A DNA "coding sequence" is a double-stranded DNA sequence which is transcribed and translated into a polypeptide *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined 5 by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic sequences, cDNA from eukaryotic mRNA, genomic DNA sequences from eukaryotic (e.g., mammalian) DNA, and even synthetic DNA sequences. A polyadenylation signal 10 and transcription termination sequence will usually be located 3' to the coding sequence.

Transcriptional and translational control sequences are DNA regulatory sequences, such as promoters, enhancers, polyadenylation signals, terminators, and the like, that provide for 15 the expression of a coding sequence in a host cell.

A "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is 20 bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of

bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site, as well as protein binding domains (consensus sequences) responsible for the binding of RNA 5 polymerase. Eukaryotic promoters often, but not always, contain "TATA" boxes and "CAT" boxes. Prokaryotic promoters typically contain Shine-Dalgarno ribosome-binding sequences in addition to the -10 and -35 consensus sequences.

An "expression control sequence" is a DNA sequence that 10 controls and regulates the transcription and translation of another DNA sequence. A coding sequence is "under the control" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then translated into the protein encoded by the coding sequence.

15 A "signal sequence" can be included near the coding sequence. This sequence encodes a signal peptide, N-terminal to the polypeptide, that communicates to the host cell to direct the polypeptide to the cell surface or secrete the polypeptide into the media, and this signal peptide is clipped off by the host cell before 20 the protein leaves the cell. Signal sequences can be found associated with a variety of proteins native to prokaryotes and eukaryotes.

As used herein, the terms "restriction endonucleases" and "restriction enzymes" refer to enzymes, each of which cut double-stranded DNA at or near a specific nucleotide sequence.

A cell has been "transformed" by exogenous or 5 heterologous DNA when such DNA has been introduced inside the cell. The transforming DNA may or may not be integrated (covalently linked) into the genome of the cell. In prokaryotes, yeast, and mammalian cells for example, the transforming DNA may be maintained on an episomal element such as a plasmid. With 10 respect to eukaryotic cells, a stably transformed cell is one in which the transforming DNA has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter 15 cells containing the transforming DNA. A "clone" is a population of cells derived from a single cell or ancestor by mitosis. A "cell line" is a clone of a primary cell that is capable of stable growth *in vitro* for many generations.

Two DNA sequences are "substantially homologous" when 20 at least about 75% (preferably at least about 80%, and most preferably at least about 90% or 95%) of the nucleotides match over

the defined length of the DNA sequences. Sequences that are substantially homologous can be identified by comparing the sequences using standard software available in sequence data banks, or in a Southern hybridization experiment under, for example, 5 stringent conditions as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, *e.g.*, Maniatis et al., *supra*; DNA Cloning, Vols. I & II, *supra*; Nucleic Acid Hybridization, *supra*.

A "heterologous" region of the DNA construct is an 10 identifiable segment of DNA within a larger DNA molecule that is not found in association with the larger molecule in nature. Thus, when the heterologous region encodes a mammalian gene, the gene will usually be flanked by DNA that does not flank the mammalian genomic DNA in the genome of the source organism. In another 15 example, coding sequence is a construct where the coding sequence itself is not found in nature (*e.g.*, a cDNA where the genomic coding sequence contains introns, or synthetic sequences having codons different than the native gene). Allelic variations or naturally-occurring mutational events do not give rise to a heterologous region 20 of DNA as defined herein.

The labels most commonly employed for these studies are radioactive elements, enzymes, chemicals which fluoresce when exposed to ultraviolet light, and others. A number of fluorescent materials are known and can be utilized as labels. These include, for 5 example, fluorescein, rhodamine, auramine, Texas Red, AMCA blue and Lucifer Yellow. A particular detecting material is anti-rabbit antibody prepared in goats and conjugated with fluorescein through an isothiocyanate. Proteins can also be labeled with a radioactive element or with an enzyme. The radioactive label can be detected by 10 any of the currently available counting procedures. The preferred isotope may be selected from  $^3\text{H}$ ,  $^{14}\text{C}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^{36}\text{Cl}$ ,  $^{51}\text{Cr}$ ,  $^{57}\text{Co}$ ,  $^{58}\text{Co}$ ,  $^{59}\text{Fe}$ ,  $^{90}\text{Y}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ , and  $^{186}\text{Re}$ . Enzyme labels are likewise useful, and can be detected by any of the presently utilized colorimetric, spectrophotometric, fluorospectrophotometric, amperometric or 15 gasometric techniques. The enzyme is conjugated to the selected particle by reaction with bridging molecules such as carbodiimides, diisocyanates, glutaraldehyde and the like. Many enzymes which can be used in these procedures are known and can be utilized. The preferred are peroxidase,  $\beta$ -glucuronidase,  $\beta$ -D-glucosidase,  $\beta$ -D- 20 galactosidase, urease, glucose oxidase plus peroxidase and alkaline phosphatase. U.S. Patent Nos. 3,654,090, 3,850,752, and 4,016,043

are referred to by way of example for their disclosure of alternate labeling material and methods.

A particular assay system developed and utilized in the art is known as a receptor assay. In a receptor assay, the material to 5 be assayed is appropriately labeled and then certain cellular test colonies are inoculated with a quantitiy of both the label after which binding studies are conducted to determine the extent to which the labeled material binds to the cell receptors. In this way, differences in affinity between materials can be ascertained.

10 An assay useful in the art is known as a "cis/trans" assay. Briefly, this assay employs two genetic constructs, one of which is typically a plasmid that continually expresses a particular receptor of interest when transfected into an appropriate cell line, and the second of which is a plasmid that expresses a reporter such as 15 luciferase, under the control of a receptor/ligand complex. Thus, for example, if it is desired to evaluate a compound as a ligand for a particular receptor, one of the plasmids would be a construct that results in expression of the receptor in the chosen cell line, while the second plasmid would possess a promoter linked to the luciferase 20 gene in which the response element to the particular receptor is inserted. If the compound under test is an agonist for the receptor,

the ligand will complex with the receptor, and the resulting complex will bind the response element and initiate transcription of the luciferase gene. The resulting chemiluminescence is then measured photometrically, and dose response curves are obtained and 5 compared to those of known ligands. The foregoing protocol is described in detail in U.S. Patent No. 4,981,784.

As used herein, the term "host" is meant to include not only prokaryotes but also eukaryotes such as yeast, plant and animal cells. A recombinant DNA molecule or gene which encodes a human 10 TADG-15 protein of the present invention can be used to transform a host using any of the techniques commonly known to those of ordinary skill in the art. Especially preferred is the use of a vector containing coding sequences for the gene which encodes a human TADG-15 protein of the present invention for purposes of prokaryote 15 transformation. Prokaryotic hosts may include *E. coli*, *S. typhimurium*, *Serratia marcescens* and *Bacillus subtilis*. Eukaryotic hosts include yeasts such as *Pichia pastoris*, mammalian cells and insect cells.

The invention includes a substantially pure DNA encoding 20 a TADG-15 protein, a DNA strand which will hybridize at high stringency to a probe containing a sequence of at least 15

consecutive nucleotides of (SEQ ID No. 1). The protein encoded by the DNA of this invention may share at least 80% sequence identity (preferably 85%, more preferably 90%, and most preferably 95%) with the amino acids listed in Figures 3 and 4 (SEQ ID No. 2). More 5 preferably, the DNA includes the coding sequence of the nucleotides of Figure 2 (SEQ ID No. 1), or a degenerate variant of such a sequence. This invention also includes a substantially pure DNA containing a sequence of at least 15 consecutive nucleotides (preferably 20, more preferably 30, even more preferably 50, and most preferably all) of 10 the region from nucleotides 1 to 3147 of the nucleotides shown in Figure 2 (SEQ ID No. 1).

By "substantially pure DNA" is meant DNA that is not part of a milieu in which the DNA naturally occurs, by virtue of separation (partial or total purification) of some or all of the molecules of that 15 milieu, or by virtue of alteration of sequences that flank the claimed DNA. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (e.g., a cDNA or a 20 genomic or cDNA fragment produced by polymerase chain reaction (PCR) or restriction endonuclease digestion) independent of other

sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence, *e.g.*, a fusion protein. Also included is a recombinant DNA which includes a portion of the nucleotides listed in Figure 2 (SEQ ID No. 1) and which 5 encodes an alternative splice variant of TADG-15.

By a "substantially pure protein" is meant a protein which has been separated from at least some of those components which naturally accompany it. Typically, the protein is substantially pure when it is at least 60% (by weight) free from the proteins and 10 other naturally-occurring organic molecules with which it is naturally associated *in vivo*. Preferably, the purity of the preparation (by weight) is at least 75%, more preferably at least 90%, and most preferably at least 99%. A substantially pure TADG-15 protein may be obtained, for example, by extraction from a natural 15 source; by expression of a recombinant nucleic acid encoding a TADG-15 polypeptide; or by chemically synthesizing the protein. Purity can be measured by any appropriate method, *e.g.*, column chromatography, such as immunoaffinity chromatography using an antibody specific for TADG-15, polyacrylamide gel electrophoresis, or 20 HPLC analysis. A protein is substantially free of naturally associated components when it is separated from at least some of those

contaminants which accompany it in its natural state. Thus, a protein which is chemically synthesized or produced in a cellular system different from the cell from which it naturally originates will be, by definition, substantially free from its naturally associated 5 components. Accordingly, substantially pure proteins include eukaryotic proteins synthesized in *E. coli*, other prokaryotes, or any other organism in which they do not naturally occur.

The term "oligonucleotide", as used herein, is defined as a molecule comprised of two or more ribonucleotides, preferably more 10 than three. Its exact size will depend upon many factors, which, in turn, depend upon the ultimate function and use of the oligonucleotide. The term "primer", as used herein, refers to an oligonucleotide, whether occurring naturally (as in a purified restriction digest) or produced synthetically, and which is capable of 15 initiating synthesis of a strand complementary to a nucleic acid when placed under appropriate conditions, *i.e.*, in the presence of nucleotides and an inducing agent, such as a DNA polymerase, and at a suitable temperature and pH. The primer may be either single-stranded or double-stranded and must be sufficiently long to prime 20 the synthesis of the desired extension product in the presence of the inducing agent. The exact length of the primer will depend upon

many factors, including temperature, sequence and/or homology of primer and the method used. For example, in diagnostic applications, the oligonucleotide primer typically contains 15-25 or more nucleotides, depending upon the complexity of the target sequence,  
5 although it may contain fewer nucleotides.

The primers herein are selected to be "substantially" complementary to particular target DNA sequences. This means that the primers must be sufficiently complementary to hybridize with their respective strands. Therefore, the primer sequence need not  
10 reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment (*i.e.*, containing a restriction site) may be attached to the 5' end of the primer, with the remainder of the primer sequence being complementary to the strand. Alternatively, non-complementary bases or longer sequences can be  
15 interspersed into the primer, provided that the primer sequence has sufficient complementary with the sequence to hybridize therewith and form the template for synthesis of the extension product.

The probe to which the DNA of the invention hybridizes preferably consists of a sequence of at least 20 consecutive  
20 nucleotides, more preferably 40 nucleotides, even more preferably 50 nucleotides, and most preferably 100 nucleotides or more (up to

100%) of the coding sequence of the nucleotides listed in Figure 2 (SEQ ID No. 1) or the complement thereof. Such a probe is useful for detecting expression of TADG-15 in a cell by a method including the steps of (a) contacting mRNA obtained from the cell with a labeled 5 TADG-15 hybridization probe; and (b) detecting hybridization of the probe with the mRNA.

By "high stringency" is meant DNA hybridization and wash conditions characterized by high temperature and low salt concentration, *e.g.*, wash conditions of 65°C at a salt concentration of 10 approximately 0.1X SSC, or the functional equivalent thereof. For example, high stringency conditions may include hybridization at about 42°C in the presence of about 50% formamide; a first wash at about 65°C with about 2X SSC containing 1% SDS; followed by a second wash at about 65°C with about 0.1X SSC.

15 The DNA may have at least about 70% sequence identity to the coding sequence of the nucleotides listed in Figure 2 (SEQ ID No. 1), preferably at least 75% (*e.g.*, at least 80%); and most preferably at least 90%. The identity between two sequences is a direct function of the number of matching or identical positions. 20 When a position in both of the two sequences is occupied by the same monomeric subunit, *e.g.*, if a given position is occupied by an

adenine in each of two DNA molecules, then they are identical at that position. For example, if 7 positions in a sequence 10 nucleotides in length are identical to the corresponding positions in a second 10-nucleotide sequence, then the two sequences have 70% sequence identity. The length of comparison sequences will generally be at least 50 nucleotides, preferably at least 60 nucleotides, more preferably at least 75 nucleotides, and most preferably 100 nucleotides. Sequence identity is typically measured using sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group (GCG), University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705).

The present invention comprises a vector comprising a DNA sequence which encodes a human TADG-15 protein, wherein said vector is capable of replication in a host, and comprises, in operable linkage: a) an origin of replication; b) a promoter; and c) a DNA sequence coding for said TADG-15 protein. Preferably, the vector of the present invention contains a portion of the DNA sequence shown in SEQ ID No. 1. Vectors may be used to amplify and/or express nucleic acid encoding a TADG-15 protein or fragment thereof.

In addition to substantially full-length proteins, the invention also includes fragments (*e.g.*, antigenic fragments) of the TADG-15 protein (SEQ ID No. 2). As used herein, "fragment," as applied to a polypeptide, will ordinarily be at least 6 residues, more 5 typically at least 9-12 residues, and preferably at least 13-20 residues in length, but less than the entire, intact sequence. Alternatively, a fragment may be an individual domain of 20-120 residues from SEQ ID No. 2. Fragments of the TADG-15 protein can be generated by methods known to those skilled in the art, *e.g.*, by 10 enzymatic digestion of naturally occurring or recombinant TADG-15 protein, by recombinant DNA techniques using an expression vector that encodes a defined fragment of TADG-15, or by chemical synthesis. The ability of a candidate fragment to exhibit a characteristic of TADG-15 (*e.g.*, binding to an antibody specific for 15 TADG-15) can be assessed by methods described herein. Purified TADG-15 or antigenic fragments of TADG-15 can be used to generate new antibodies or to test existing antibodies (*e.g.*, as positive controls in a diagnostic assay) by employing standard protocols known to those skilled in the art. Included in this invention is polyclonal 20 antisera generated by using TADG-15 or a fragment of TADG-15 as the immunogen in, *e.g.*, rabbits. Standard protocols for monoclonal

and polyclonal antibody production known to those skilled in this art are employed. The monoclonal antibodies generated by this procedure can be screened for the ability to identify recombinant TADG-15 cDNA clones, and to distinguish them from other cDNA 5 clones.

Further included in this invention are TADG-15 proteins which are encoded, at least in part, by portions of SEQ ID No. 2, *e.g.*, products of alternative mRNA splicing or alternative protein processing events, or in which a section of TADG-15 sequence has 10 been deleted. The fragment, or the intact TADG-15 polypeptide, may be covalently linked to another polypeptide, *e.g.*, one which acts as a label, a ligand or a means to increase antigenicity.

The invention also includes a polyclonal or monoclonal antibody which specifically binds to TADG-15. The invention 15 encompasses not only an intact monoclonal antibody, but also an immunologically-active antibody fragment, *e.g.*, a Fab or (Fab)<sub>2</sub> fragment; an engineered single chain Fv molecule; or a chimeric molecule, *e.g.*, an antibody which contains the binding specificity of one antibody, *e.g.*, of murine origin, and the remaining portions of 20 another antibody, *e.g.*, of human origin.

In one embodiment, the antibody, or a fragment thereof, may be linked to a toxin or to a detectable label, *e.g.*, a radioactive label, non-radioactive isotopic label, fluorescent label, chemiluminescent label, paramagnetic label, enzyme label, or colorimetric label. Examples of suitable toxins include diphtheria toxin, *Pseudomonas* exotoxin A, ricin, and cholera toxin. Examples of suitable enzyme labels include malate hydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, alcohol dehydrogenase, alpha-glycerol phosphate dehydrogenase, triose phosphate isomerase, peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase, acetylcholinesterase, etc. Examples of suitable radioisotopic labels include  $^3\text{H}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^{14}\text{C}$ , etc.

Paramagnetic isotopes for purposes of *in vivo* diagnosis can also be used according to the methods of this invention. There are numerous examples of elements that are useful in magnetic resonance imaging. For discussions on *in vivo* nuclear magnetic resonance imaging, see, for example, Schaefer et al., (1989) *JACC* 14, 472-480; Shreve et al., (1986) *Magn. Reson. Med.* 3, 336-340; Wolf, G. L., (1984) *Physiol. Chem. Phys. Med. NMR* 16, 93-95; Wesbey et al.,

(1984) *Physiol. Chem. Phys. Med. NMR* 16, 145-155; Runge et al., (1984) *Invest. Radiol.* 19, 408-415. Examples of suitable fluorescent labels include a fluorescein label, an isothiocyalate label, a rhodamine label, a phycoerythrin label, a phycocyanin label, an 5 allophycocyanin label, an ophthaldehyde label, a fluorescamine label, etc. Examples of chemiluminescent labels include a luminal label, an isoluminal label, an aromatic acridinium ester label, an imidazole label, an acridinium salt label, an oxalate ester label, a luciferin label, a luciferase label, an aequorin label, etc.

10 Those of ordinary skill in the art will know of other suitable labels which may be employed in accordance with the present invention. The binding of these labels to antibodies or fragments thereof can be accomplished using standard techniques commonly known and used by those of ordinary skill in the art.

15 Typical techniques are described by Kennedy et al., (1976) *Clin. Chim. Acta* 70, 1-31; and Schurs et al., (1977) *Clin. Chim. Acta* 81, 1 -

40. Coupling techniques mentioned in the latter are the glutaraldehyde method, the periodate method, the dimaleimide method, the m-maleimidobenzyl-N-hydroxy-succinimide ester 20 method. All of these methods are incorporated by reference herein.

Also within the invention is a method of detecting TADG-15 protein in a biological sample, which includes the steps of contacting the sample with the labeled antibody, *e.g.*, radioactively tagged antibody specific for TADG-15, and determining whether the antibody binds to a component of the sample. Antibodies to the TADG-15 protein can be used in an immunoassay to detect increased levels of TADG-15 protein expression in tissues suspected of neoplastic transformation. These same uses can be achieved with Northern blot assays and analyses.

As described herein, the invention provides a number of diagnostic advantages and uses. For example, the TADG-15 protein is useful in diagnosing cancer in different tissues since this protein is highly overexpressed in tumor cells. Antibodies (or antigen-binding fragments thereof) which bind to an epitope specific for TADG-15, are useful in a method of detecting TADG-15 protein in a biological sample for diagnosis of cancerous or neoplastic transformation. This method includes the steps of obtaining a biological sample (*e.g.*, cells, blood, plasma, tissue, etc.) from a patient suspected of having cancer, contacting the sample with a labeled antibody (*e.g.*, radioactively tagged antibody) specific for TADG-15, and detecting the TADG-15 protein using standard immunoassay techniques such as an ELISA.

Antibody binding to the biological sample indicates that the sample contains a component which specifically binds to an epitope within TADG-15.

Likewise, a standard Northern blot assay can be used to 5 ascertain the relative amounts of TADG-15 mRNA in a cell or tissue obtained from a patient suspected of having cancer, in accordance with conventional Northern hybridization techniques known to those of ordinary skill in the art. This Northern assay uses a hybridization probe, *e.g.*, radiolabelled TADG-15 cDNA, either containing the full-length, single stranded DNA having a sequence complementary to 10 SEQ ID No. 1 (Figure 2), or a fragment of that DNA sequence at least 20 (preferably at least 30, more preferably at least 50, and most preferably at least 100 consecutive nucleotides in length). The DNA hybridization probe can be labeled by any of the many different 15 methods known to those skilled in this art.

The following examples are given for the purpose of illustrating various embodiments of the invention and are not meant to limit the present invention in any fashion.

## EXAMPLE 1

### Tissue collection and storage

Upon patient hysterectomy, bilateral salpingo-  
5 oophorectomy, or surgical removal of neoplastic tissue, the specimen  
is retrieved and placed on ice. The specimen was then taken to the  
resident pathologist for isolation and identification of specific tissue  
samples. Finally, the sample was frozen in liquid nitrogen, logged  
into the laboratory record and stored at -80°C.

10 Additional specimens were frequently obtained from the  
Cooperative Human Tissue Network (CHTN). These samples were  
prepared by the CHTN and shipped on dry ice. Upon arrival, these  
specimens (e.g., blood (serum), urine, saliva, tears and insterstitial  
fluid) were logged into the laboratory record and stored at -80°C.  
15 Participation of the following divisions of the Cooperative Human  
Tissue Network (CHTN) in providing tumor tissues is acknowledged:  
Western Division, Case Western Reserve University, (Cleveland, OH);  
Midwestern Division, Ohio state University, (Columbus, OH); Eastern  
Division, NDRI, (Philadelphia, PA); Pediatric Division, Children's  
20 Hospital, (Columbus, OH); Southern Division, University of Alabama at  
Birmingham, (Birmingham, AL).

primers according to the manufacturer's protocol utilizing a first strand synthesis kit obtained from CLONTECH (Palo Alto, CA). The purity of the cDNA was evaluated by PCR using primers specific for the p53 gene. These primers span an intron such that pure cDNA can 5 be distinguished from cDNA that is contaminated with genomic DNA.

### EXAMPLE 3

10 PCR with redundant primers, cloning of TADG-15 cDNA, T-vector ligation and transformations and DNA sequencing

Redundant primers,

forward 5'-TGGGTIGTIACIGCIGCICA(C/T)TG-3' (SEQ ID No. 11) and reverse 5'-A(A/G)IGGICCICCI(C/G)(T/A)(A/G)TCICC-3' (SEQ ID 15 No. 12), corresponding to the amino acids surrounding the catalytic triad for serine proteases, were used to compare the PCR products from normal and carcinoma cDNAs.

The purified PCR products were ligated into the Promega T-vector plasmid and the ligation products used to transform JM109 20 competent cells according to the manufacturer's instructions (Promega). Positive colonies were cultured for amplification, the

III/grade III ovarian adenocarcinoma patient. Three overlapping clones were obtained which spanned 3147 nucleotides.

5

#### EXAMPLE 4

##### Northern blot analysis

10  $\mu$ g mRNAs were size separated by electrophoresis through a 1% formaldehyde-agarose gel in 0.02 M MOPS, 0.05 M sodium acetate (pH 7.0), and 0.001 M EDTA. The mRNAs were then blotted to Hybond-N<sup>+</sup> nylon membrane (Amersham) by capillary action in 20x SSPE. The RNAs are fixed to the membrane by baking for 2 hours at 80°C.  $^{32}$ P-labeled cDNA probes were made by Prime-a-Gene Labeling System (Promega). The PCR products amplified by 15 the same primers described above were used for probes. The blots were prehybridized for 30 min and hybridized for 60 min at 68°C with  $^{32}$ P-labeled cDNA probe in ExpressHyb Hybridization Solution (CLONTECH). Control hybridization to determine relative gel loading was performed with a  $\beta$ -tubulin probe.

Normal human tissues; spleen, thymus, prostate, testis, ovary, small intestine, colon and peripheral blood leukocyte, and normal human fetal tissues; brain, lung, liver and kidney (Human Multiple Tissue Northern Blot; CLONTECH) were also examined by the 5 same hybridization procedure. Additional multiple tissue northern (MTN) blots from CLONTECH include the Human MTN blot, the Human MTN II blot, the Human Fetal MTN II blot, and the Human Brain MTN III blot.

10

#### EXAMPLE 5

##### Western blot analysis

Polyclonal rabbit antibody was generated by 15 immunization with a poly-lysine linked multiple Ag peptide derived from the TADG-15 protein sequence 'LFRDWIKENTGV' (SEQ ID No. 13). Approximately 20 µg of cell lysates were separated on a 15% SDS-PAGE gel and electroblotted to PVDF at 100 V for 40 min at 4°C. The proteins were fixed to the membrane by incubation in 50% 20 MeOH for 10 min. The membrane was blocked overnight in TBS (pH 7.8) containing 0.2% non-fat milk. Primary antibody was added to

the membrane at a dilution of 1:100 in 0.2% milk/TBS and incubated for 2 h at room temperature. The blot was washed and incubated with a 1:3000 dilution of alkaline-phosphatase conjugated goat anti-rabbit IgG (BioRad) for 1 h at room temperature. The blot was 5 washed and incubated with a chemiluminescent substrate before a 10 sec exposure to X-ray film for visualization.

#### EXAMPLE 6

10

##### Quantitative PCR

The mRNA overexpression of TADG-15 was determined using a quantitative PCR. Quantitative PCR was performed.<sup>11,12</sup> Oligonucleotide primers were used 15 for TADG-15:

forward 5'-ATGACAGAGGATTCAAGGTAC-3' (SEQ ID No. 14) and

reverse 5'-GAAGGTGAAGTCATTGAAGA-3' (SEQ ID No. 15); and

and for  $\beta$ -tubulin:

forward 5'-CGCATCAACGTGTACTACAA-3' (SEQ ID No. 16) and

20 reverse 5'-TACGAGCTGGTGGACTGAGA-3' (SEQ ID No. 17).

$\beta$ -tubulin was utilized as an internal control.

The PCR reaction mixture consists of cDNA derived from 50 ng of mRNA, 5 pmol of sense and antisense primers for both the TADG-15 gene and the  $\beta$ -tubulin gene, 200  $\mu$ mol of dNTPs, 5  $\mu$ Ci of  $\alpha$ - $^{32}$ PdCTP and 0.25 units of Taq DNA polymerase with reaction buffer 5 (Promega) in a final volume of 25  $\mu$ l. The target sequences were amplified in parallel with the  $\beta$ -tubulin gene. Thirty cycles of PCR were carried out in a Thermal Cycler (Perkin Elmer Gene Amp 2400; Perkin-Elmer Cetus). Each cycle of PCR included 30 sec of denaturation at 94°C, 30 sec of annealing at 60°C and 30 sec of 10 extension at 72°C. The annealing temperature varies according to the primers that are used in the PCR reaction. For the reactions involving degenerate primers, an annealing temperature of 48°C was used. The appropriate annealing temperature for the TADG-15- and  $\beta$ -tubulin-specific primers is 62°C.

15 A portion of the PCR products were separated on 2% agarose gels and the radioactivity of each PCR product was determined by using a PhosphoImager (Molecular Dynamics). In the present study, the expression ratio (TADG-15/ $\beta$ -tubulin) was used to evaluate gene expression and defined the value at mean  $\pm$  2SD of 20 normal ovary as the cut-off value to determine overexpression. The

student's *t* test was used for comparison of the mean values of normal ovary and tumors.

5

### EXAMPLE 7

#### Immunohistochemistry

Immunohistochemical staining was performed using a Vectastain Elite ABC Kit (Vector). Formalin-fixed and paraffin-10 embedded specimens were routinely deparaffinized and processed using microwave heat treatment in 0.01 M sodium citrate buffer (pH 6.0). The specimens were incubated with normal goat serum in a moist chamber for 30 min. After incubation with biotinylated anti-rabbit IgG for 30 min, the sections were then incubated with ABC 15 reagent (Vector) for 30 min. The final products were visualized using the AEC substrate system (DAKO) and sections were counterstained with hematoxylin before mounting. Negative controls were performed using normal serum instead of the primary antibody.

20

## EXAMPLE 8

### Antisense TADG-15

TADG-15 is cloned and expressed in the opposite orientation such that an antisense RNA molecule (SEQ ID No. 18) is produced. For example, the antisense RNA is used to hybridize to the complementary RNA in the cell and thereby inhibit translation of TADG-15 RNA into protein.

10

## EXAMPLE 9

### Peptide ranking

For vaccine or immune stimulation, individual 9-mers to 11-mers were examined to rank the binding of individual peptides to 15 the top 8 haplotypes in the general population (Parker et al., (1994)). The computer program used for this analyses can be found at [http://www-bimas.dcrt.nih.gov/molbio/hla\\_bind/](http://www-bimas.dcrt.nih.gov/molbio/hla_bind/). Table 1 shows the peptide ranking based upon the predicted half-life of each peptide's binding to a particular HLA allele. A larger half-life 20 indicates a stronger association with that peptide and the particular HLA molecule. The TADG-15 peptides that strongly bind to an HLA

allele are putative immunogens, and are used to inoculate an individual against TADG-15.

5

**TABLE 1**

TADG-15 peptide ranking

|           | <u>HLA Type &amp; Ranking</u> | <u>Start</u> | <u>Peptide</u> | <u>Predicted Dissociation<sub>1/2</sub></u> | <u>SEQ ID No.</u> |
|-----------|-------------------------------|--------------|----------------|---|-------------------|
| HLA A0201 |                               |              |                |   |                   |
| 10        | 1                             | 68           | VLLGIGFLV      | 2537.396                                    | 19                |
|           | 2                             | 126          | LLYSGVPFL      | 1470.075                                    | 20                |
| 15        | 3                             | 644          | SLISPWNLV      | 521.640                                     | 21                |
|           | 4                             | 379          | KVSFKFFYL      | 396.525                                     | 22                |
|           | 5                             | 386          | YLLEPGVPA      | 346.677                                     | 23                |
|           | 6                             | 257          | SLTFRSFDL      | 123.902                                     | 24                |
|           | 7                             | 762          | ILQKGEIRV      | 118.238                                     | 25                |
| 20        | 8                             | 841          | RLPLFRDWI      | 106.842                                     | 26                |
|           | 9                             | 64           | GLLLVLLGI      | 88.783                                      | 27                |
|           | 10                            | 57           | VLAAVLIGL      | 83.527                                      | 28                |
| HLA A0205 |                               |              |                |   |                   |
| 25        | 1                             | 67           | LVLLGIGFL      | 142.800                                     | 29                |
|           | 2                             | 379          | KVSFKFFYL      | 100.800                                     | 30                |
|           | 3                             | 126          | LLYSGVPFL      | 71.400                                      | 31                |
|           | 4                             | 88           | KVFNGYMRI      | 36.000                                      | 32                |
|           | 5                             | 670          | TQWTAFLGL      | 33.600                                      | 33                |
|           | 6                             | 119          | KVKDALKLL      | 25.200                                      | 34                |
| 30        | 7                             | 60           | AVLIGLLL       | 24.000                                      | 35                |
|           | 8                             | 62           | LIGLLLVLL      | 23.800                                      | 36                |
|           | 9                             | 57           | VLAAVLIGL      | 23.800                                      | 37                |
|           | 10                            | 61           | VLIGLLLVL      | 23.800                                      | 38                |
| HLA A1    |                               |              |                |   |                   |
| 35        | 1                             | 146          | FSEGSVIAY      | 337.500                                     | 39                |
|           | 2                             | 658          | YIDDRGFRY      | 125.000                                     | 40                |
|           | 3                             | 449          | SSDPCPGQF      | 75.000                                      | 41                |

|    |         |     |           |         |
|----|---------|-----|-----------|---------|
|    |         |     | (         | )       |
|    |         |     |           |         |
|    | 4       | 401 | YVEINGEKY | 45.000  |
|    | 5       | 387 | LLEPGVPAG | 18.000  |
|    | 6       | 553 | GSDEASCPK | 15.000  |
|    | 7       | 97  | TNENFVDAY | 11.250  |
| 5  | 8       | 110 | STEFVSLAS | 11.250  |
|    | 9       | 811 | SVEADGRIF | 9.000   |
|    | 10      | 666 | YSDPTQWTA | 7.500   |
|    | HLA A24 |     |           |         |
| 10 | 1       | 709 | DYDIALLEL | 220.000 |
|    | 2       | 408 | KYCGERSQF | 200.000 |
|    | 3       | 754 | QYGGTGALI | 50.000  |
|    | 4       | 153 | AYYWSEFSI | 50.000  |
|    | 5       | 722 | EYSSMVRPI | 50.000  |
|    | 6       | 326 | GFEATFFQL | 36.000  |
| 15 | 7       | 304 | TFHSSQNVL | 24.000  |
|    | 8       | 707 | TFDYDIALL | 20.000  |
|    | 9       | 21  | KYNSRHEKV | 16.500  |
|    | 10      | 665 | RYSDPTQWT | 14.400  |
|    | HLA B7  |     |           |         |
| 20 | 1       | 686 | APGVQERRL | 240.000 |
|    | 2       | 12  | GPKDFGAGL | 80.000  |
|    | 3       | 668 | DPTQWTAFL | 80.000  |
|    | 4       | 461 | TGRCIRKEL | 60.000  |
|    | 5       | 59  | AAVLIGLLL | 36.000  |
| 25 | 6       | 379 | KVSFKFFYL | 20.000  |
|    | 7       | 119 | KVKDALKLL | 20.000  |
|    | 8       | 780 | LPQQITPRM | 20.000  |
|    | 9       | 67  | LVLLGIGFL | 20.000  |
|    | 10      | 283 | SPMEPHALV | 18.000  |
| 30 | HLA B8  |     |           |         |
|    | 1       | 12  | GPKDFGAGL | 24.000  |
|    | 2       | 257 | SLTFRSFDL | 8.000   |
|    | 3       | 180 | MLPPRARSL | 8.000   |
|    | 4       | 217 | GLHARGVEL | 8.000   |
| 35 | 5       | 173 | MAEERVVML | 4.800   |
|    | 6       | 267 | SCDERGSDL | 4.800   |
|    | 7       | 567 | CTKHTYRCL | 4.000   |
|    | 8       | 724 | SSMVRPICL | 4.000   |
|    | 9       | 409 | YCGERSQFV | 3.600   |
| 40 | 10      | 495 | TCKNKFCKP | 3.200   |

|    |           |     |           |          |    |
|----|-----------|-----|-----------|----------|----|
|    | HLA B2702 |     |           |          |    |
| 5  | 1         | 427 | VRFHSDQSY | 1000.000 | 79 |
|    | 2         | 695 | KRIISHPFF | 600.000  | 80 |
|    | 3         | 664 | FRYSDPTQW | 500.000  | 81 |
|    | 4         | 220 | ARGVELMRF | 200.000  | 82 |
|    | 5         | 492 | HQFTCKNKF | 100.000  | 83 |
|    | 6         | 53  | GRWVVLAAC | 100.000  | 84 |
|    | 7         | 248 | LRGDADSVL | 60.000   | 85 |
|    | 8         | 572 | YRCLNGLCL | 60.000   | 86 |
|    | 9         | 692 | RRLKRIISH | 60.000   | 87 |
|    | 10        | 24  | SRHEKVNGL | 60.000   | 88 |
|    | HLA B4403 |     |           |          |    |
| 15 | 1         | 147 | SEGSVIAYY | 360.000  | 89 |
|    | 2         | 715 | LELEKPAEY | 360.000  | 90 |
|    | 3         | 105 | YENSNSTEF | 60.000   | 91 |
|    | 4         | 14  | KDFGAGLKY | 50.625   | 92 |
|    | 5         | 129 | SGVPFLGPY | 36.000   | 93 |
|    | 6         | 436 | TDTGFLAEY | 33.750   | 94 |
|    | 7         | 766 | GEIRVINQT | 30.000   | 95 |
|    | 8         | 402 | VEINGEKYC | 30.000   | 96 |
|    | 9         | 482 | DELNCSCDA | 24.000   | 97 |
|    | 10        | 82  | RDVRVQKVF | 22.500   | 98 |

25

### EXAMPLE 10

#### TADG-15 cDNA

30 A screening strategy to identify proteases which are overexpressed in human cancer has been developed in which RT-PCR products amplified specifically in tumors, as compared to normal

( ( tissue, are examined.<sup>9</sup> During this effort, candidate genes were identified using redundant sense primers to the conserved amino acid histidine domain at the NH<sub>2</sub> end of the catalytic domain and antisense primers to the downstream conserved amino acid serine 5 domain. Subcloning and sequencing the appropriate 480 base pair band(s) amplified in such a PCR reaction provides the basis for identifying the gene(s) encoding proteases(s). Among these amplified catalytic domains, a new serine protease gene named TADG-15 (tumor antigen-derived gene 15) was identified. The 10 catalytic domain of the newly identified TADG-15 protein is similar to other serine proteases and specifically contains conserved amino acids appropriate for the catalytic domain of the trypsin-like serine protease family.

A computerized search of GenEMBL databases using the 15 FASTA program (Wisconsin Package Version 9.1, GCG, Madison, Wisconsin) for amino acid sequences homologous to the TADG-15 protease domain revealed that homologies with other known human proteases never exceeds 55%. Figure 1 shows the alignment of the protease domain of TADG-15 compared with other human serine 20 proteases. Using the BESTFIT program available through GCG, the

similarities between TADG-15 and trypsin, chymotrypsin, and tissue-type plasminogen activator are 51%, 46% and 52%, respectively.

From the sequence derived from the TADG-15 catalytic domain, specific primers were synthesized to amplify a TADG-15-specific probe for library screening. After screening an ovarian carcinoma library, one 1785 bp clone was obtained which included the 3' end of the TADG-15 transcript. Upon further screening using the 5' end of the newly detected clone, two additional clones were identified which provided another 1362 bp of the cDNA, including the 5' end of the TADG-15 transcript. The total length of the sequenced cDNA was approximately 3.15 kb. The total nucleotide sequence obtained includes a Kozak's consensus sequence preceding a single open reading frame encoding a predicted protein of 855 amino acids (Figure 2).

The deduced open reading frame encoded by the TADG-15 nucleotide sequence (Figures 2, 3 and 4) contains several distinct domains as follows: an amino terminal cytoplasmic tail (amino acids (aa) #1-54), a potential transmembrane domain (aa #55-77), an extracellular membrane domain (aa #78-213), two complement subcomponents Clr/Cls, Uegf, and bone morphogenetic protein 1 (CUB) repeats (aa #214-447), four ligand binding repeats of the low

density lipoprotein (LDL) receptor-like domain (aa #453-602) and a serine protease domain (aa #615-855). The TADG-15 protein also contains two potential N-linked glycosylation sites (aa #109 and 302) and a potential proteolytic cleavage site upstream from the protease domain (aa #614) which could release and/or activate the protease at the carboxy end of this protein. In addition, TADG-15 contains an RGD motif (aa #249-251) which is commonly found in proteins involved in cell-cell adhesion.

10

#### EXAMPLE 11

##### TADG-15 expression

To examine the size of the transcript for TADG-15 and its pattern of expression in various tissues, Northern blot hybridization was performed for representative histological types of carcinoma and in a series of cell lines, fetal tissues and normal adult tissues (Figure 5). The transcript size for the TADG-15 message was determined to be approximately 3.2 kb and a single intense transcript appeared to be present in all of the carcinomas examined, whereas no visible band was detected in normal ovary (Figure 5). This transcript size is

also in good agreement with the sequence data predicting a transcript size of 3.15 kb. The ovarian tumor cell lines, SW626 and CAOV3, also showed an abundance of transcript, however little or no transcript was detectable in the breast carcinoma cell lines MDA-MB-5 231 and MDA-MB-4355. Among normal human fetal tissues, fetal kidney showed an abundance of the TADG-15 transcript and low expression was also detected in fetal lung. In normal adult tissues, TADG-15 was detected in colon with low levels of expression in small intestine and prostate (Figure 5).

10 To evaluate mRNA transcript expression of TADG-15 in ovarian tumors and normal ovary, semi-quantitative PCR (Figure 6) was performed. In a preliminary study, the linearity of this assay<sup>11,12</sup> was confirmed and its efficacy correlated with both Northern blots and immunohistochemistry. The data was quantified 15 using a phosphoimager and compared as a ratio of expression (TADG-15/β-tubulin). Results herein indicate that TADG-15 transcript expression is elevated above the cut-off value (mean for normal ovary  $\pm$  2 SD) in all of the tumor cases examined and is either not detected or detected at extremely low levels in normal ovaries 20 (Figure 6A and B). Analysis of ovarian carcinoma subtypes, including early stage and late stage disease, confirms overexpression of TADG-

15 in all carcinomas examined (Table 2). All of the carcinomas studied, which included 5 stage I and 3 stage II carcinomas, showed overexpression of the TADG-15 gene.

These data can also be examined with regard to tumor 5 stage and histological sub-type, and results indicated that every carcinoma of every stage and histological sub-type overexpressed the TADG-15 gene. The expression ratio (mean value  $\pm$  SD) for normal ovary group was determined as  $0.182 \pm 0.024$ , for LMP tumor group as  $0.847 \pm 0.419$  and for carcinoma group as  $0.771 \pm 10 0.380$  (Table 2). A comparison between the normal ovary group and tumor groups showed that overexpression of the TADG-15 gene is statistically significant in both the LMP tumor group and the carcinoma group (LMP tumor:  $p<0.001$ , carcinoma:  $p<0.0001$ ).

As shown in Figure 6, TADG-15 transcripts were noted in 15 all ovarian carcinomas, but were not present at detectable levels in any of the following tissues: a) normal ovary, b) fetal liver and brain, c) adult spleen, thymus, testes, ovary and peripheral blood lymphocytes, d) skeletal muscle, liver, brain or heart. This evaluation was extended to a standard panel of about 40 tumors. 20 Using TADG-15-specific primers, the expression was also examined in

tumor cell lines derived from both ovarian and breast carcinoma tissues as shown in Figure 7 and in other tumor tissues as shown in Figure 8. Expression of TADG-15 was also observed in carcinomas of the breast, colon, prostate and lung.

5 Polyclonal antibodies developed to a synthetic peptide (a 12-mer) at the carboxy terminus of the protease domain were used to examine TADG-15 expression in cell lines by Western blot and by immunolocalization in normal ovary and ovarian tumors. Western blots of cell extracts from SW626 and CAOV3 cells were probed with  
10 both antibody and preimmune sera (Figure 9). Several bands were detected with the antibody, including bands of approximately 100,000 daltons, approximately 60,000 daltons and 32,000 daltons. The anticipated molecular size of the complete TADG-15 molecule is estimated to be approximately 100,000 daltons, and the protease  
15 domain which may be released by proteolytic cleavage at aa #614 is estimated to be approximately 32,000 daltons. Some intermediate proteolytic product may be represented by the 60,000 dalton band.

Antibody staining of tumor cells confirms the presence of the TADG-15 protease in the cytoplasm of a serous LMP tumor, 20 mucinous LMP tumor and serous carcinoma (Figure 10B, C & D, respectively). This diffuse staining pattern may be due to detection

of TADG15 within the cell as it is being packaged and transported to the cell surface. In endometrioid carcinoma, the antigen is clearly detectable on the surface of tumor cells (Figure 10E). No staining was detected in normal ovarian epithelium or stromal cells (Figure 5 10A). Immunohistochemical staining of a series of 27 tumors indicates the presence of the TADG-15 protein in all the carcinoma subtypes examined, including the low malignant potential group. Strong staining was noted in 7 of 9 low malignant potential tumors and 13 of 18 carcinomas (Table 3).

10

**TABLE 2**

Number of cases with overexpression of TADG-15 in normal ovaries and ovarian tumors

|              | <u>N</u> | <u>overexpression</u> | <u>expression ratio<sup>a</sup></u> |
|--------------|----------|-----------------------|-------------------------------------|
|              |          | <u>of TADG-15</u>     |                                     |
| 5            |          |                       |                                     |
| Normal       | 10       | 0 (0%)                | 0.182 ± 0.024                       |
| LMP          | 10       | 10 (100%)             | 0.847 ± 0.419                       |
| 10           |          |                       |                                     |
| serous       | 6        | 6 (100%)              | 0.862 ± 0.419                       |
| mucinous     | 4        | 4 (100%)              | 0.825 ± 0.483                       |
| Carcinoma    | 31       | 31 (100%)             | 0.771 ± 0.380                       |
| 15           |          |                       |                                     |
| serous       | 18       | 18 (100%)             | 0.779 ± 0.332                       |
| mucinous     | 7        | 7 (100%)              | 0.907 ± 0.584                       |
| endometrioid | 3        | 3 (100%)              | 0.502 ± 0.083                       |
| clear cell   | 3        | 3 (100%)              | 0.672 ± 0.077                       |

<sup>a</sup>The ratio of expression level of TADG-15 to  $\beta$ -tubulin (mean ± SD)

**TABLE 3**

Immunohistochemical staining using TADG-15

|    | <u>Lab No.</u> | <u>Histology</u>                | <u>TADG-15</u> |
|----|----------------|---------------------------------|----------------|
| 5  |                | Surface epithelium of the ovary | -              |
|    | H-3194         | serous (LMP)                    | ++             |
|    | H-162          | serous (LMP)                    | ++             |
|    | H-1182         | serous (LMP)                    | ++             |
| 10 | H-4818         | serous (LMP)                    | ++             |
|    | H-4881         | serous (LMP)                    | ++             |
|    | H-675          | mucinous (LMP)                  | +              |
|    | H-2446         | mucinous (LMP)                  | +              |
|    | H-0707         | mucinous (LMP)                  | ++             |
| 15 | H-2042         | mucinous (LMP)                  | ++             |
|    | H-2555         | serous carcinoma                | ++             |
|    | H-1858         | serous carcinoma                | ++             |
|    | H-5266         | serous carcinoma                | ++             |
|    | H-5316         | serous carcinoma                | +              |
| 20 | H-2597         | serous carcinoma                | +              |
|    | H-4931         | mucinous carcinoma              | ++             |
|    | H-1867         | mucinous carcinoma              | ++             |
|    | H-5998         | mucinous carcinoma              | ++             |
|    | H-2679         | endometrioid adenocarcinoma     | +              |
| 25 | H-5718         | endometrioid adenocarcinoma     | ++             |
|    | H-3993         | endometrioid adenocarcinoma     | +              |
|    | H-2991         | endometrioid adenocarcinoma     | ++             |
|    | H-2489         | endometrioid adenocarcinoma     | ++             |
|    | H-5994         | clear cell carcinoma            | ++             |
| 30 | H-6718         | clear cell carcinoma            | ++             |
|    | H-1661         | clear cell carcinoma            | ++             |
|    | H-6201         | clear cell carcinoma            | ++             |
|    | H-5640         | clear cell carcinoma            | +              |

- Negative; + Weak Positive; ++ Strong Positive (more than 50% of cell staining)

## EXAMPLE 12

### TADG-15 homology

Recently, a mouse protein named epithin (GenBank Accession No. AF042822) has been described.<sup>14</sup> Epithin is a 902 amino acid protein which contains a similar structure to TADG-15 in that it has a cytoplasmic domain, transmembrane domain, two CUB domains, four LDLR-like domains and a carboxy terminal serine protease domain. TADG-15 and epithin are 84% similar over 843 amino acids, suggesting that the proteins may be orthologous (Figure 11). The precise role of epithin remains to be elucidated.

A search of GeneBank for similar previously identified sequences yielded one such sequence with relatively high homology to a portion of the TADG-15 gene. The similarity between the portion of TADG-15 from nucleotide #182 to 3139 and SNC-19 (GeneBank Accession No. #U20428) is approximately 97% (Figure 12). There are however significant differences between SNC-19 and TADG-15. For example, TADG-15 has an open reading frame of 855 amino acids whereas the longest open reading frame of SNC-19 is 173 amino acids. Additionally, SNC-19 does not include a proper start site for the initiation of translation, nor does it include the

amino terminal portion of the protein encoded by TADG-15. Moreover, SNC-19 does not include an open reading frame for a functional serine protease because the His, Asp and Ser residues of the catalytic triad that are necessary for function are encoded in 5 different reading frames.

### Implications

The overall structure of the TADG-15 protein is relatively similar to the members of the tolloid/BMP-1 family and the 10 complement subcomponents, Clr/Cls. These proteins contain both CUB and protease domains, and complex formation through the ligand binding domain is essential for their function. Activation of the serine protease domains of Clr and Cls requires proteolytic cleavage of Arg-Gly and Arg-Ile bonds, respectively.<sup>15</sup> Similarly, it 15 might be expected that the TADG-15 protein is synthesized as a zymogen, which is activated by cleavage between Arg<sup>614</sup> and Val<sup>615</sup> and analogous to the activation mechanism of other serine protease zymogens. Western blot analysis of cultured cell lysates confirmed both a 100 kDa and 32 kDa peptide, which correspond to the putative 20 zymogen (whole molecule) and a cleaved protease product of TADG-15 (Figure 9). These data support a model for proteolytic release

and/or activation of TADG-15 as occurs for similar type II serine proteases.

CUB domains were first found in complement subcomponents C1r/C1s<sup>16-18</sup> and are known to be a widespread module in developmentally regulated proteins, such as the bone morphogenetic protein-1 (BMP-1) and the tolloid gene product.<sup>18-20</sup> The role of these repeats remains largely unknown. However, some models suggest that the CUB domain may be involved in protein-protein interactions. The CUB domain of Clr and Cls participates in the assembly of the Cls-Clr-Clr-Cls tetrameric complex in the activation of the classical pathway of complement by providing protein-protein interaction domains.<sup>15</sup> The *Drosophila* decapentaplegic (DPP) protein is essential for dorsal-ventral specification of the embryo, and the *Drosophila* tolloid (TLD) forms a complex with DPP to regulate its activity.<sup>19,20</sup> Missense mutations in the CUB domain of the tolloid protein results in a phenotype that does not allow a protein interaction with the DPP complex.<sup>19</sup>

The TADG-15 protein contains two tandem repeats of CUB-like domains between amino acid residues 214 and 447. Each of these is approximately 110 amino acids long and each has four conserved cysteine residues characteristic of other CUBs (amino acids

214, 244, 268, 294, 340, 366, 397, 410). By analogy, the CUB repeats of the TADG-15 protein may form an interactive domain capable of promoting multimeric complex formation and regulating the activity of the target protein or TADG-15 itself.

5                   The TADG-15 protein also contains the LDL receptor ligand binding repeat (class A motif) -like domain, which consists of four contiguous cysteine-rich repeats (amino acid residues 453 to 602). Each cysteine-rich repeat is approximately 40 amino acids long and contains a conserved, negatively-charged sequence (Ser-Asp-10 Glu) with six cysteine residues. In the LDL receptor protein, this repeat is thought to function as a protein-binding domain which interacts with the lysine and arginine residues present in lipoproteins.<sup>21,22</sup> In addition, the first repeat of the LDL receptor appears to bind  $\text{Ca}^{2+}$  and not the lipoproteins.<sup>23</sup> By analogy, it is 15 possible that the LDL receptor-like repeat in TADG-15 may act in a similar fashion, interacting with positively charged regions of other proteins and/or as a  $\text{Ca}^{2+}$  binding site. As a result of ligand binding and the formation of receptor-ligand complex, LDL receptor is internalized via clathrin-coated pits.<sup>24</sup> These types of plasma 20 membrane receptors contain a characteristic amino acid sequence in their cytoplasmic domain for binding to clathrin-coated pits.<sup>24</sup> TADG-

15 does not contain this motif in its cytosolic region, and furthermore, no similarities with other known protein sequences were found in the cytoplasmic domain of the TADG-15. This finding suggests that TADG-15 functions in a different manner from the 5 endocytic receptors (such as the LDL receptor), although TADG-15 possesses similar ligand-binding repeats in the extracellular matrix.

Although the precise role of TADG-15 is unknown, this gene is clearly overexpressed in ovarian tumors. A variety of proteases, such as type IV collagenase and plasminogen activator, 10 appear to be involved in the process of tumor invasion and are constituents of a protease cascade in malignant progression. TADG-15 may constitute such an activity and directly digest extracellular matrix components surrounding a tumor, or activate other proteases by cleavage of inactive precursors, indirectly enhancing tumor 15 growth and invasion. It is also possible that TADG-15 may function like a member of the tolloid/BMP-1 family by forming complexes with other growth factors or signal transduction proteins to modulate their activities.

These data raise the possibility that the TADG-15 gene 20 and its translated protein will be a useful marker for the early detection of ovarian carcinoma through release of the protease

domain into the extracellular matrix and ultimately the circulation. These data also suggest the possibility of using TADG-15 as a target for therapeutic intervention through delivery systems directed at the CUB/LDLR ligand binding domains.

5                   The following references were cited herein:

1.    Liotta, L.A., et al. *Cell*, **64**: 327-336, 1991.
2.    Duffy, M.J. *Clin. Exp. Metastasis*, **10**: 145-155, 1992.
3.    Tryggvason, K., et al. *Biochem. Biophys. Acta.*, **907**: 191-217, 1987.
- 10   4.    Levy, A.T., et al. *Cancer Res.*, **51**: 439-444, 1991.
5.    Monsky, W.L. et al. *Cancer Res.*, **53**: 3159-3164, 1993.
6.    Duffy, M.J. et al. *Cancer*, **62**: 531-533, 1988.
7.    Häckel, C., et al. *Cancer*, **79**: 53-58, 1997.
8.    Watt, K, et al. *Proc. Natl. Acad. Sci. U.S.A.*, **83**: 3166-3170, 1986.
- 15   9.    Tanimoto, H. et al. *Cancer Res.*, **57**: 2884-2887, 1997.
11.   Shigemasa, K. et al. *J. Soc. Gynecol. Invest.*, **4**: 95-102, 1997.
12.   Tanimoto, H. et al. *Gynecol. Oncol.*, **66**: 308-312, 1997.
13.   Maniatis, T., Fritsch, E.F. & Sambrook, J. *Molecular Cloning*, p. 309-361 Cold Spring Harbor Laboratory, New York, 1982.
- 20   14.   Kim, M.G., et al. *Immunogenetics*, **49**(5): 420-428, 1999.
15.   Arlaud et al. *Method in Enzymology*, **223**: 61-82, 1993.

16. Journet, A. & Tosi, M. *Biochem. J.*, 240: 783-787, 1986.

17. Mackinnon, C.M., et al. *Eur. J. Biochem.*, 169: 547-553, 1987.

18. Bork, P. & Beckmann, G. *J. Mol. Biol.*, 231: 539-545, 1993.

19. Childs, S.R. & O'Connor, M.B. *Dev. Biol.*, 162: 209-220, 1994.

5 20. Blader, PL, et al. *Science*, 278: 1937-1940, 1997.

21. Yamamoto, T. et al. *Cell*, 39: 27-38, 1984.

22. Daly, N.L., et al. *Proc. Natl. Acad. Sci.*, 92: 6334-6338, 1995.

23. van Driel, I.R., et al. *J. Biol. Chem.*, 262: 17443-17449, 1987.

24. Lodish, H. et al. *Sorting of membrane proteins internalized from*  
10 *the cell surface. In: Molecular Cell Biology*, 3rd ed., p.722-733  
*Scientific American Books, Inc., New York, 1995.*

25. Parker, KC et al. *J. Immunol.* 152:163, 1994.

Any patents or publications mentioned in this  
15 specification are indicative of the levels of those skilled in the art to  
which the invention pertains. These patents and publications are  
herein incorporated by reference to the same extent as if each  
individual publication was specifically and individually indicated to  
be incorporated by reference.

20 One skilled in the art will readily appreciate that the  
present invention is well adapted to carry out the objects and obtain

the ends and advantages mentioned, as well as those inherent therein. The present examples along with the methods, procedures, treatments, molecules, and specific compounds described herein are presently representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art which are encompassed within the spirit of the invention as defined by the scope of the claims.